

Drone based Mapping and Identification of Young Spruce Stand for Semiautonomous Cleaning^{*}

Issouf Ouattara^{*} Heikki Hyyti^{**} Arto Visala^{*}

^{*} *Department of Electrical Engineering and Automation,
Aalto University, Espoo, Finland*

^{**} *Finnish Geospatial Research Institute (FGI),
National Land Survey of Finland*

Abstract: We propose a novel method to locate spruces in a young stand with a low cost unmanned aerial vehicle. The method has three stages: 1) the forest area is mapped and a digital surface model and terrain models are generated, 2) the locations of trees are found from a canopy height model using local maximum and watershed algorithms, and 3) these locations are used in a convolutional neural network architecture to detect young spruces. Our result for detecting young spruce trees among other vegetation using only color images from a single RGB camera were promising. The proposed method is able to achieve a detection accuracy of more than 91%. As low cost unmanned aerial vehicles with color cameras are versatile today, the proposed work is enabling low cost forest inventory for automating forest management.

Keywords: unmanned aerial vehicle, mapping, individual tree identification, convolutional neural network, autonomous vehicle, forestry.

1. INTRODUCTION

Unmanned aerial vehicles are increasingly used in many application areas and particularly in agriculture and forestry. The data collected by drones can enable automated precision forestry and lead to accurate and efficient decision making. This is particularly important in semiautonomous or autonomous forest management operations such as in cleaning which consists of removing weeds and competing deciduous trees around young planted spruces. These forest management operations are required to help the young spruces to thrive on a planted forest stand. This work has been inefficient to mechanize (see e.g. Hallongren and Rantala 2013; Saksä et al. 2016) so more autonomous solutions would be required (Vestlund and Hellström, 2006).

According to the Natural Resource Institute of Finland, seventy five percent (75%) of the land area is covered by

^{*} Strategic Research Council at the Academy of Finland is acknowledged for financial support of project “Competence-Based Growth Through Integrated Disruptive Technologies of 3D Digitalization, Robotics, Geospatial Information and Image Processing/Computing - Point Cloud Ecosystem” (project decision numbers 293389 and 314312).

Issouf Ouattara is a PhD Student with the Autonomous Systems research group, Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, PO Box 15500, FIN-00076 Aalto, Finland (e-mail: issouf.ouattara@aalto.fi).

Heikki Hyyti is a researcher with the department of Remote Sensing and Photogrammetry at the Finnish Geospatial Research Institute (FGI), National Land Survey of Finland (e-mail: heikki.hyyti@iki.fi).

Arto Visala is a Professor with the Autonomous Systems research group, Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, PO Box 15500, FIN-00076 Aalto, Finland (e-mail: arto.visala@aalto.fi).

forest making the country the most extensive forest cover country in Europe. Forest stands are mainly privately owned and the size of the average forest holding is about 25 hectares (Holopainen et al., 2014). According to Heinonen et al. (2017), the Finnish forest industry may need yearly 10–30 million m^3 more wood in the coming years. This increasing need in production coupled with the need of increased profitability of forestry require new approaches in data collection and analysis for forest management tasks such as cleaning, thinning, and logging. The information requirement for each phase of forests’ maturity varies.

In this study we are interested in detecting young spruce in seedling forest stands (young spruce forest stands) with spruces higher than one meter and which have not been cleaned on time. In those conditions, deciduous trees and other undesirable species grow faster than the young spruces and might cover/occlude the young spruces. The task for manual or autonomous cleaning by machine becomes challenging because the spruces might not be visible by the driver or the sensors mounted on the machine. Automated solutions such as the method for detecting young spruces using machine vision developed by Vihlman et al. (2015) have limitations as they might not be able to detect occluded spruce trees. Hyyti et al. (2013) developed a solution that used images from a camera mounted on point cleaning tool of the cleaning machine. When the tool is lifted up, the camera can capture images from above and a detection algorithm is used to detect young spruces in the image. However, this setup is limited by how high the point cleaning tool can be lifted above from the ground and it is not suitable for an environment where young spruces and other trees are several meters high. This is the limitation of

ground vehicles that have mounted camera or laser scanner setups. In this situation, aerial vehicle mounted sensors are preferred. An aerial image from an unmanned aerial vehicle (UAV) is more suitable for seedling forest stands for ground truth inventory (Holopainen et al., 2014). UAVs can acquire high resolution images which makes them suitable for seedling forest stands inventory. Several studies have attempted to use UAV based imagery or a lidar platform to detect and identify trees from above. The majority of these studies tries to solve two problems: individual tree identification (ITD) and the classification of the identified trees.

Individual tree identification:

Nevalainen et al. (2017) uses RGB images and hyperspectral images in the tree detection and classification process. In their work, RGB images are used to generate point cloud data which then are used to generate an orthophotomap. For the classification task, point cloud features combined with hyperspectral features were used. They achieved good detection results using Random forest and Multilayer Perceptron. Lidar data has also been used for individual tree detection. This is done by either using directly the lidar point cloud data (Zhang et al., 2015), or by using the lidar point cloud to derive a canopy height model (CHM) from which the trees are identified using local maxima algorithm (Popescu and Wynne, 2004; Katoh et al., 2017). Hyperspectral sensors and decent lidar sensors can be expensive and are not readily available and accessible. In addition, for quadcopter UAV system, the weight of payload directly affects the flight time of the system. Usage of only an RGB camera can increase the flight time of the system, and therefore, impact the overall cost of the data acquisition process. Several studies have investigated the use of UAV-acquired RGB images and the derived data such as point clouds, digital terrain model (DTM), digital surface model (DSM), and canopy height model (CHM) for individual tree identification (Baena et al., 2017; Huang et al., 2018; Mohan et al., 2017). In this case, either the derived canopy model is used (Baena et al., 2017), or the derived RGB orthophoto is directly used (Huang et al., 2018).

Classification of trees:

In the application of UAV data in forest study, tree classification generally follows the individual tree identification step. The goal of this step is to differentiate one or more tree species from the rest of the image content. Two methodologies seems to emerge for tree classification in UAV images. The first uses selected features (or variables) which can be extracted from the UAV dataset (Kampen et al., 2019; Dash et al., 2017; Puliti et al., 2018). These variables include spectral variables, textural variables, and geometric variables. The second approach uses directly deep learning for tree classification (Onishi and Ise, 2018; Natesan et al., 2019). In both cases the trees are first segmented, and each segmented tree or its features are passed to the classifier. Both methods heavily depend on the accuracy of the segmentation step.

The objective of this study is to develop a young spruce identification method using only RGB image data from a UAV in a young forest stand which has not been cleaned on time. Previous related studies have used CHM for tree identification and a classification algorithm has been used

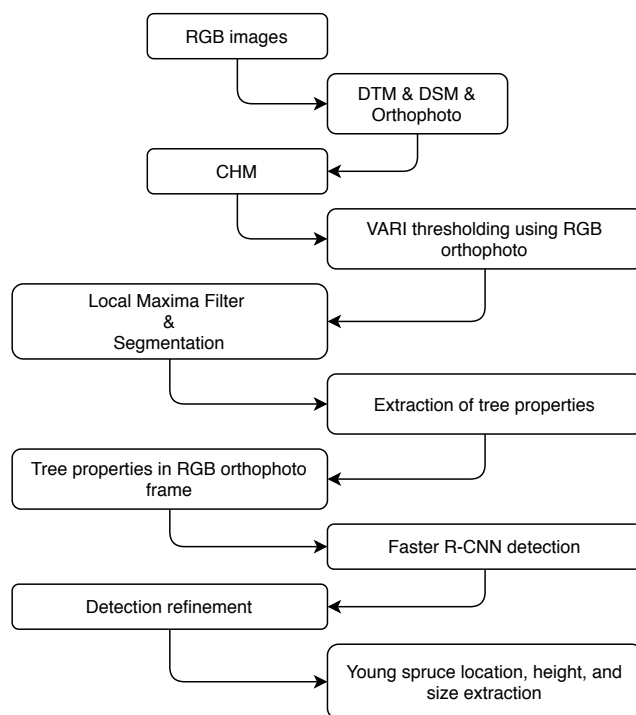


Fig. 1. The proposed process of young spruce identification and classification

in the detection and classification process. This study uses a convolutional neural network (CNN) based detection algorithm following a local maxima detection step from CHM raster image. The advantage of our approach is that the classification step is not tightly linked to the segmentation performance. In fact the classification step in this study is used to enhance the segmentation step. The following paper is structured so that the materials and methods are presented in Section 2, the results are given in Section 3 and discussed in Section 4, and finally conclusions are drawn in Section 5.

2. MATERIALS AND METHODS

This section introduces the materials and methods used in this study. Figure 1 shows the overall process of the proposed method. In the method, first, RGB images are collected with an unmanned aerial vehicle. These images are used to build a DTM, a DSM, and an orthophotomap. A CHM is derived from the DTM and DSM raster data. The orthophotomap is used to compute a vegetation index which is used to remove non vegetation data from the CHM. The resulting CHM is used to segment potential trees using local maxima filter followed by a watershed segmentation. Tree properties are extracted from the segmentation and translated to the RGB image frame. Finally the RGB orthophoto along with the tree properties is used to detect spruces. The detection is done using the Faster R-CNN inception v2 followed by a refinement for removing non valid trees. The following sections explain the parts (shown in Figure 1) of the proposed method in detail.

2.1 Images Collection and Processing

In this study, nadir RGB images collected with a UAV are used for tree identification and detection. The DJI

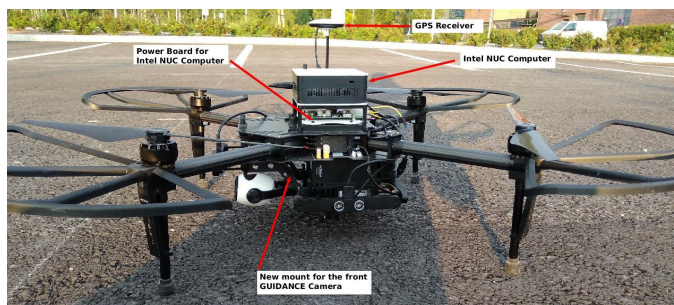


Fig. 2. Our DJI Matrice 100 drone with RGB camera

Matrice 100 UAV shown in Figure 2 has been used in this study. It is a quadcopter with a 650 mm diagonal length. It uses N1 flight controller with the latest firmware which is proprietary. The aircraft has also several sensors for attitude control, position holding and navigation purposes. It is equipped with the Zenmuse Z3 camera used to take the nadir images at specific waypoints locations. The waypoints locations have been chosen such that the images have about 70% horizontal overlaps in x and y directions. A total of 680 images were collected using this system on the 10th of August 2019 and 11th of September 2019. Additionally, 200 images have been collected on 22nd of October 2019 when deciduous trees started losing their leaves. Each image collection campaign covers only part of the study area. This is important because the data will be split into training data and validation data in Section 2.3.

The collected images are used to produce an orthophoto, a digital surface model (DSM), and digital terrain model (DTM) using the Agisoft Metashape Professional v1.5.5 software (Agisoft LLC, 2019). The DSM is generated by directly using the dense point cloud data derived from the collected images. In order to generate the DTM, the point cloud is segmented in two sets: ground and non ground point cloud. The ground point cloud is used to generate the DTM. These data are generated using the reference projection frame WGS84/UTM zone 35N (EPSG::32635). The DTM and the DSM have a resolution of 5.4 cm/pixel while the orthophoto has a resolution of 2.5 cm/pixel .

2.2 Individual Tree Identification (ITD)

All the subsequent data processing in this study were performed using python language and libraries. The goal was to develop a complete method using one language. Python is used because of rapid prototyping capability and the richness of the python environment.

A canopy height model (CHM) is generated by subtracting the DTM from the DSM. The CHM represents the height of any object above ground. In order to remove low vegetation (e.g., bushes) from the data, the CHM has been filtered with a threshold value of 0.5 meter. Anything below 0.5 meter is then discriminated. This value was chosen based on the young spruce height which is above one meter in the study area. Additional step is also used to remove non vegetation objects such as man made objects and people from the data. This is achieved with Visible Atmospherically Resistant Index (VARI) (Gitelson et al., 2002).

The CHM is used to identify individual trees using a local maxima filter. Variable window size based on tree height have been used in a previous study (Popescu and Wynne, 2004) in the local maxima filter. However, their model used to infer the width of the tree from the height data has been derived using only limited number of data points and does not apply necessarily to the tree data in this study. Therefore, a simpler approach using fixed window size such as in the study by Mohan et al. (2017) is preferred. We used a fixed window of 5×5 pixels. The local maxima filter results in a set of pixel locations representing the tops of trees. These tree tops are used as markers in a marker controlled Watershed algorithm which segments the trees patches using the CHM raster data (Kornilov and Safonov, 2018). The scikit-image morphology *watershed* function and the *regionprops* function have been used to segment tree patches and retrieve statistical information such as *equivalent diameter*, *coordinate list of tree patch*, *major axis length*, and *minor axis length* (van der Walt et al., 2014). These statistical parameters and a principal component analysis (PCA) have been used to discard invalid segmentations. Each tree segment is represented by a set of connected 2D coordinate points, i.e., list of pixels corresponding to the segment. In order to measure how circular the segmented tree patch is, the PCA is run over the 2D coordinates representing the segmented tree to fit an ellipse to the 2D coordinate points. The PCA returns two eigenvectors and corresponding eigenvalues which characterise the distribution of the points in a plane. The ratio of the two eigenvalues and the value of the highest eigenvalue are used, by setting threshold values, to discard segmented regions whose shape is too small or elongated.

2.3 Young Spruce Tree Detection

The high resolution RGB orthophoto is used for the young spruce detection. The CHM resolution is almost two times lower than the orthophoto resolution. For detection/classification this high resolution orthophoto is required, and therefore, downsampling to the orthophoto resolution needs to be avoided. In order to overlay the identified trees tops and patches from the CHM to the orthophoto, a frame transformation is required. This study takes advantage of the fact that both images (CHM and orthophoto) have reference projected frame attached to them. Using coordinate transformation, both the tree tops locations and the tree patches are translated from the CHM frame to the orthophoto frame. Each image patch can be used as an input for classification. This however heavily depends on the segmentation process. If the region segmented as a tree contains both a spruce and a deciduous tree, it might be classified as a deciduous tree if the spruce only occupies a small portion of the image. A sliding window classifier could be used on the orthophoto map as suggested by Guirado et al. (2017). This, however, does not take advantage of the individual tree identification (ITD) step.

In this study Faster R-CNN with Inception v2 architecture is used in conjunction with the individual tree identification step for young spruce detection. Faster R-CNN is a deep learning detection algorithm that uses region proposal network (RPN) for rapid prediction of object bounds

and objectness scores (Ren et al., 2015) in contrast with its predecessor Fast R-CNN which used time consuming selective search for region proposal (Girshick, 2015). The Inception v2 network is used as the feature extractor by the Faster R-CNN for extracting feature maps from input image (Szegedy et al., 2015). The advantage of using the R-CNN over the classification techniques is two-fold. When an image patch contains both spruces and broad-leaved trees, it can detect the spruce among other trees and avoid the bulky classification error. It avoids the need for sliding window classification which involves choosing the window size of classification and also the sliding rate. Faster R-CNN is already finding the potential locations of objects of interest using a region proposal network (RPN) (Ren et al., 2015).

Patches of trees from the ITD step were annotated and used for training the Faster R-CNN network. Transfer learning has been used which reduces the need for a large dataset during training. It tries to transfer the knowledge learned from one task to a new related task thereby improving the learning in the new task (Gorban and Zinovyev, 2009). A pre-trained Faster R-CNN inception v2 model from the Tensorflow model zoo is used and trained using the Tensorflow object detection API (Huang et al., 2016). A total of 3000 images (200×200 pixels) were used for the training. These images have been retrieved from the orthophoto built using the collected RGB images by the drone (cf. Section 2.1) and using basic image augmentation techniques such as rotation and flipping which exist in the Tensorflow object detection API (Huang et al., 2016). About 20% of the data was manually picked out for testing. The annotation is done using the annotation software Labelling by Tzutalin (2015).

A final refinement step is used after the output of Faster R-CNN Inception v2 output. It tries to remove detections that are not representative of a spruce tree. These detections may, for example, be a small part of a spruce tree. They are removed using principal component analysis (PCA) on the pixel coordinates of each detection. This is done by using the 2D coordinate points contained in the bounding box of detected trees. These 2D points are used as inputs to the PCA algorithm resulting in two eigenvalues. By thresholding the ratio of the two eigenvalues and the value of the highest eigenvalue, detections that are not representative of spruce tree are removed.

2.4 Study Area

The study area was a young spruce stand at Lohja, a city and municipality in the Uusimaa region of Finland. The site is located at approximately $60^{\circ}25'18.8''N$ and $24^{\circ}05'54.7''E$. It has an approximate area of 8 acres. The study area is shown on Figure 3. The young spruce stand in this area has not been cleaned on time making it a good study area for this project. The average spruce height is above one meter. The deciduous trees have grown up and make it impossible for a camera mounted on a forest machine to see all of the young spruce trees. The objective is therefore to map this young spruce stand from above and identify the young spruce trees which cannot otherwise be located using ground vehicle mounted camera. To evaluate the method developed in this study, two subareas have



Fig. 3. Study area ($60^{\circ}25'18.8''N$ and $24^{\circ}05'54.7''E$)

been chosen inside the study area. The number of young spruce trees inside these subareas have been manually counted and will serve as a reference.

3. RESULTS

The previous section presented the methods and materials used for identifying young spruces and classifying them. In the first step, trees were identified in the study area using a CHM (derived from the DTM and DSM) and a fixed window local maxima filter. A watershed segmentation algorithm is used with the tree tops to segment the tree patches. In the second step, images of tree patches are passed through a CNN in order to find young spruces inside each patch.

Individual Tree Identification:

The output of the individual tree identification is a set of points representing tree tops. This set of points is translated from the CHM frame to the RGB orthophoto frame. Figure 4 shows the result of ITD. As visible in the figure, most of the young spruces are identified in this step. For deciduous trees, because they don't have a conic shape, the number of found tops does not reflect the real number of tree tops. However, this is not a problem since the main objective is to detect spruce trees.

When a spruce tree is surrounded by taller broad-leaved trees, the spruce might not be detected in this step because its top is not the local maxima. This could be resolved by reducing the maxima filter window size. However reducing the window size results in over-segmentation in the segmentation phase. Therefore, the window size is kept at 5×5 pixels.

The set of tree tops is used in the marker controlled watershed algorithm in order to segment out tree patches. The result is shown in Figure 5.

Tree Classification and Detection refinement:

The ITD was able to find trees using local maximum filter on the CHM raster data and segment the tree patches using marker-controlled watershed algorithm. These tree patches can be passed through a classification algorithm to find out if a given patch is a spruce or deciduous tree (Onishi and Ise, 2018; Natesan et al., 2019). However, the classification has a shortcoming as it is not able to classify

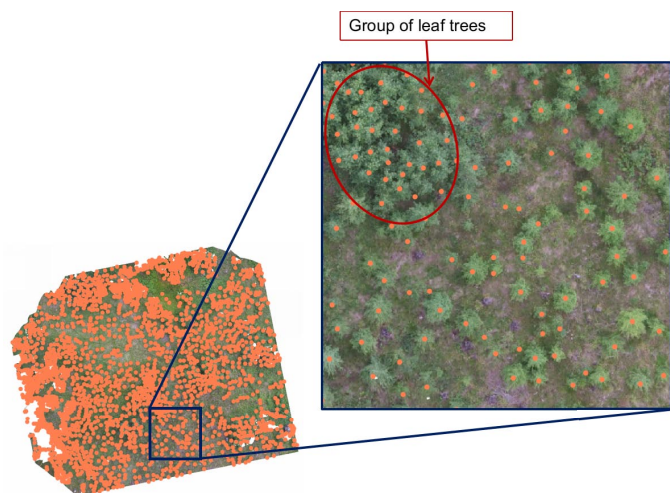


Fig. 4. Identified tree tops

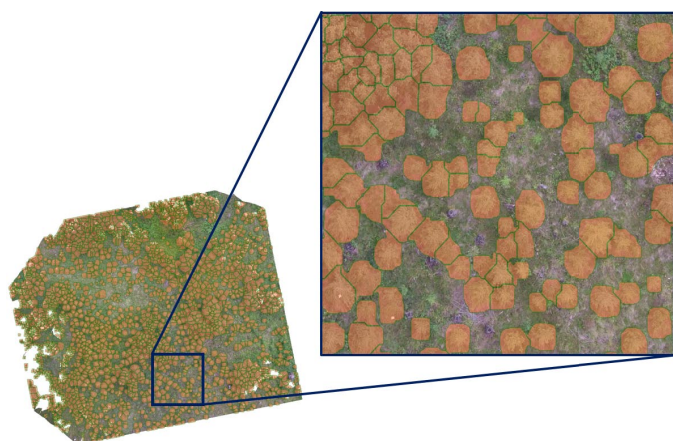


Fig. 5. Segmented tree patches

correctly a patch as a spruce if the patch contains a spruce and a larger portion of it is occupied by deciduous trees. In order to overcome this shortcoming, the patches are passed through a detection algorithm instead of a classification algorithm. The difference of these is that the detection algorithm tries to find which part of the image is a spruce tree, while a classification algorithm tries to find if the overall input image is a spruce or not.

The result of the detection is shown in Figure 6. Detections that are either a small segment of spruce tree or the form is not representative of a spruce tree are automatically found and removed using principle component analysis. Figure 7 shows the result of this refinement for sample trees.

Overall young spruce detection error:

In the identification step, the identification error has not been evaluated as it represents only an intermediate step in this study and the number of deciduous trees identified is difficult to measure because they are very dense. The overall detection error represents the error in detecting young spruces. The error is evaluated for two areas in two periods of time. The first period is during summer time and the second period is during autumn when deciduous trees change color and loose their leaves.

The detection and classification errors are measured using overall detection accuracy, omission error and commission

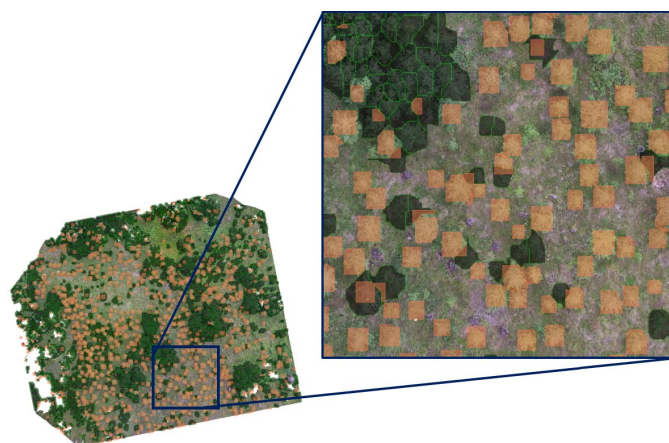


Fig. 6. Detection result



Fig. 7. Refined detection result

error. The overall detection accuracy is the ratio of the number of correctly detected spruces by the reference number of spruces (shorthand # spruces in Table 1). The commission error is defined as the ratio of false positives to all positives, and the omission error is defined as the ratio of false negatives to the amount of targets (i.e., the sum of false negatives and true positives) (Nuijten et al., 2019).

The detection accuracy for the **area A** is 96.04% during summer and 91.96% during autumn. The commission and omission error for **area A** are respectively 10.18%, and 3.96% during summer and 2.86%, and 8.03% during autumn. For **area B**, the detection accuracy is 94.11% for summer data and 95.27% for autumn data. The commission and omission error for **area B** are respectively 12.5%, and 5.88% during summer and 6.20%, and 4.72% during autumn.

Table 1. Error analysis: detection accuracy, commission error, and omission error (A_1, B_1 : area A, and B during summer, A_2, B_2 : area A, and B during autumn).

Area	A_1	B_1	A_2	B_2
# spruces	101	119	112	127
true positives	97	112	103	121
false positives	11	16	3	8
false negatives	4	7	9	6
detection (%)	96.04	94.11	91.96	95.27
commission (%)	10.18	12.5	2.86	6.20
omission (%)	3.96	5.88	8.03	4.72

Visual inspection shows that the commission error is mainly a misclassification error, meaning an error due to the CNN algorithm falsely classifies other things as spruce trees. While the omission error is due to both the errors during the tree top detection step and the classification step. The omission error occurs in the tree top detection step when the local maxima algorithm failed to find the top of the tree. During the detection step, the omission error occurs when the Faster R-CNN failed to detect a spruce inside a segment from the ITD step.

4. DISCUSSION

In this study, we proposed a method for detecting young spruce trees in a seedling stand. The locations of the detected young spruce trees can be used in an autonomous or semi-autonomous forest machine for automatic cleaning. Our results show that the method developed in this study is effective in both periods of time (summer and autumn), reaching more than 91% detection accuracy. This is important because there are forest sites in Finland that cannot be operated in certain period of time. Developing young spruce detection method that is robust is necessary to get timely and accurate data whenever needed. However, the method have only been tested in a small area. Subsequent tests will be needed to validate these results, especially with higher amount of manually counted reference trees.

The detection method is a combination of an individual tree detection (ITD) step and a convolutional neural network (CNN) based detection which is used to improve the result of the ITD step and classify the young spruces. This is intuitively correct because the ITD step might give a tree segment that contains both spruce and potentially other trees. Thus limiting to the use of classification only might lead to an error. In this study, it has not been evaluated how much the CNN improve the ITD step. This could be evaluated in the future work to give more rigorous understanding of the usage of CNN based detection algorithm instead of a simple classifier.

Since the goal of the study was to provide accurate positions of young spruce trees for aiding automatic cleaning, the positions were reported in WGS 84 coordinate system. However, the overall positioning accuracy of the satellite navigation system was not evaluated in this study and should be done in the future work. A better satellite navigation solution can also be used to ensure that the orthophoto map has a valid location information. This could be achieved with the help of Real-Time Kinematic (RTK) or Post Processed Kinematic (PPK) positioning method during image data collection as shown by Tomaščík et al. (2019).

5. CONCLUSION

In this work we have used a low-cost drone and its camera to collect nadir RGB images from above a young spruce stand to detect young spruce trees for automated forest management operations. The proposed method is able to detect most individual spruce trees in spite that they might be occluded by other faster growing vegetation such as competing deciduous trees. The method effectively combines local maxima on a canopy height model (CHM),

watershed segmentation, and a trained Faster R-CNN inception v2 model deep neural network.

Forest management operations are required to help the young spruces to thrive. This work has been challenging to automate earlier using only forest machine mounted sensors. Now on the contrary, we have shown that we are able to detect at least 91% of young spruces among other vegetation. This enables the development of more automated forest management practices in the future, where the semiautonomous forest machine may guide its operator to effectively cut competing trees around young spruce seedlings.

REFERENCES

- Agisoft LLC, St. Petersburg, R. (2019). *Agisoft Metashape User Manual, Professional Edition, Version 1.5*. URL https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf.
- Baena, S., Moat, J., Whaley, O., and Boyd, D.S. (2017). Identifying species from the air: UAVs and the very high resolution challenge for plant conservation. *PLOS ONE*, 12(11), 1–21. doi:10.1371/journal.pone.0188714.
- Dash, J., Pearce, G., Watt, M., and Paul, T. (2017). Combining Airborne Laser Scanning and Aerial Imagery Enhances Echo Classification for Invasive Conifer Detection. *Remote Sensing*, 9(2), 156. doi:10.3390/rs9020156.
- Girshick, R.B. (2015). Fast R-CNN. *CoRR*, abs/1504.08083. URL <http://arxiv.org/abs/1504.08083>.
- Gitelson, A.A., Kaufman, Y.J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 80(1), 76–87. doi:10.1016/S0034-4257(01)00289-9.
- Gorban, A.N. and Zinovyev, A.Y. (2009). Handbook of research on machine learning applications and trends : Algorithms, methods, and techniques. 242–264.
- Guirado, E., Tabik, S., Alcaraz-Segura, D., Cabello, J., and Herrera, F. (2017). Deep-learning convolutional neural networks for scattered shrub detection with Google Earth imagery. *CoRR*, abs/1706.00917. URL <http://arxiv.org/abs/1706.00917>.
- Hallongren, H. and Rantala, J. (2013). A search for better competitiveness in mechanized early cleaning through product development: evaluation of two naarva uprooters. *International Journal of Forest Engineering*, 24(2), 91–100.
- Heinonen, T., Pukkala, T., Mehtätalo, L., Asikainen, A., Kangas, J., and Peltola, H. (2017). Scenario analyses for the effects of harvesting intensity on development of forest resources, timber supply, carbon balance and biodiversity of finnish forestry. *Forest Policy and Economics*, 80, 80 – 98. doi:<https://doi.org/10.1016/j.forpol.2017.03.011>.
- Holopainen, M., Vastaranta, M., and Hyyppä, J. (2014). Outlook for the next generation's precision forestry in Finland. *Forests*, 5(7), 1682–1694. doi:10.3390/f5071682.
- Huang, H., Li, X., and Chen, C. (2018). Individual tree crown detection and delineation from very-high-resolution UAV images based on bias field and marker-controlled watershed segmentation algorithms. *IEEE Journal of Selected Topics in Applied Earth Observa-*

- tions and Remote Sensing*, 11(7), 2253–2262. doi:10.1109/JSTARS.2018.2830410.
- Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., and Murphy, K. (2016). Speed/accuracy trade-offs for modern convolutional object detectors. *CoRR*, abs/1611.10012. URL <http://arxiv.org/abs/1611.10012>.
- Hyyti, H., Kalmari, J., and Visala, A. (2013). Real-time detection of young spruce using color and texture features on an autonomous forest machine. In *the 2013 International Joint Conference on Neural Networks (IJCNN)*, 1–8. doi:10.1109/IJCNN.2013.6707122.
- Kampen, M., Lederbauer, S., Ján, Mund, P.D., and Immitzer, M. (2019). UAV-based multispectral data for tree species classification and tree vitality analysis.
- Katoh, M., Deng, S., Takenaka, Y., Cheung, K., Oono, K., Horisawa, M., Hyyppä, J., Yu, X., Liang, X., and Wang, Y. (2017). Development of smart precision forest in conifer plantation in Japan using laser scanning data. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3/W3, 95–100. doi:10.5194/isprs-archives-XLII-3-W3-95-2017.
- Kornilov, A.S. and Safonov, I.V. (2018). An overview of watershed algorithm implementations in open source libraries. *J. Imaging*, 4, 123.
- Mohan, M., Silva, C.A., Klauberg, C., Jat, P., Catts, G., Cardil, A., Hudak, A.T., and Dia, M. (2017). Individual tree detection from unmanned aerial vehicle (UAV) derived canopy height model in an open canopy mixed conifer forest. *Forests*, 8(9). URL <https://www.mdpi.com/1999-4907/8/9/340>.
- Natesan, S., Armenakis, C., and Vepakomma, U. (2019). Resnet-Based Tree Species Classification Using UAV Images. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4213, 475–481. doi:10.5194/isprs-archives-XLII-2-W13-475-2019.
- Nevalainen, O., Honkavaara, E., Tuominen, S., Viljanen, N., Hakala, T., Yu, X., Hyyppä, J., Saari, H., Pölonen, I., Imai, N.N., and Tommaselli, A.M.G. (2017). Individual tree detection and classification with UAV-based photogrammetric point clouds and hyperspectral imaging. *Remote Sensing*, 9(3). doi:10.3390/rs9030185.
- Nuijten, R.J., Coops, N.C., Goodbody, T.R., and Pelletier, G. (2019). Examining the multi-seasonal consistency of individual tree segmentation on deciduous stands using digital aerial photogrammetry (DAP) and unmanned aerial systems (UAS). *Remote Sensing*, 11(7). doi:10.3390/rs11070739.
- Onishi, M. and Ise, T. (2018). Automatic classification of trees using a UAV onboard camera and deep learning. *CoRR*, abs/1804.10390. URL <http://arxiv.org/abs/1804.10390>.
- Popescu, S.C. and Wynne, R.H. (2004). Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height.
- Puliti, S., Talbot, B., and Astrup, R. (2018). Tree-stump detection, segmentation, classification, and measurement using unmanned aerial vehicle (UAV) imagery.
- Ren, S., He, K., Girshick, R.B., and Sun, J. (2015). Faster R-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497. URL <http://arxiv.org/abs/1506.01497>.
- Saksa, T. et al. (2016). Proceedings of the OSCAR workshop: Mechanized and efficient silviculture: November 25–26, 2015 Natural Resources Institute Finland, Suonenjoki Research Unit, Finland.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567. URL <http://arxiv.org/abs/1512.00567>.
- Tomaščík, J., Mokroš, M., Surový, P., Grznárová, A., and Merganič, J. (2019). UAV RTK/PPK method—an optimal solution for mapping inaccessible forested areas? *Remote Sensing*, 11(6). doi:10.3390/rs11060721.
- Tzutalin (2015). *LabelImg. Git code (2015)*. URL <https://github.com/tzutalin/labelImg>.
- van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., and the scikit-image contributors (2014). scikit-image: image processing in Python. *PeerJ*, 2, e453. doi:10.7717/peerj.453.
- Vestlund, K. and Hellström, T. (2006). Requirements and system design for a robot performing selective cleaning in young forest stands. *Journal of terramechanics*, 43(4), 505–525.
- Vihlman, M., Hyyti, H., Kalmari, J., and Visala, A. (2015). Detection and species classification of young trees using machine perception for a semi-autonomous forest machine. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 1543–1548. doi:10.1109/ICRA.2015.7139394.
- Zhang, C., Zhou, Y., and Qiu, F. (2015). Individual tree segmentation from lidar point clouds for urban forest inventory. *Remote Sensing*, 7(6), 7892–7913. doi:10.3390/rs70607892.